Stochastic Game の開放型 Tree Queuing Network への応用(最適化数理の手法と実際)

毛利, 裕昭

数理解析研究所講究録 1461 巻 2005 年 214-224

Stochastic Game の開放型 Tree Queuing Network への応用

早稲田大学・商学部　理工学部　数学・応用数学研究所
毛利裕昭(Hiroaki Mohri)
School of Commerce, Institute of Pure and Applied Mathematics
Waseda University
mohri@waseda.jp

1．はじめに

ここでは，グラフが木構造をもった計算機ネットワークの制御に確率過程ゲーム（Stochastic Game）を適用することを考える．確率過程ゲームについては，Mertens(2002), Vieille(2002), Raghavan and Filar(1991)がサーベイ論文としては便利である．関連書籍としては，Filar and Vrieze (1996)がある．

Mohri(2005)では,有限個の並列待ち行列のスイッチングの問題についてAltman(1996)に沿った考え方に基づき，モデルを記述して定常状態および有限ステージでの最適方策（ゲームの戦略）について考察した．

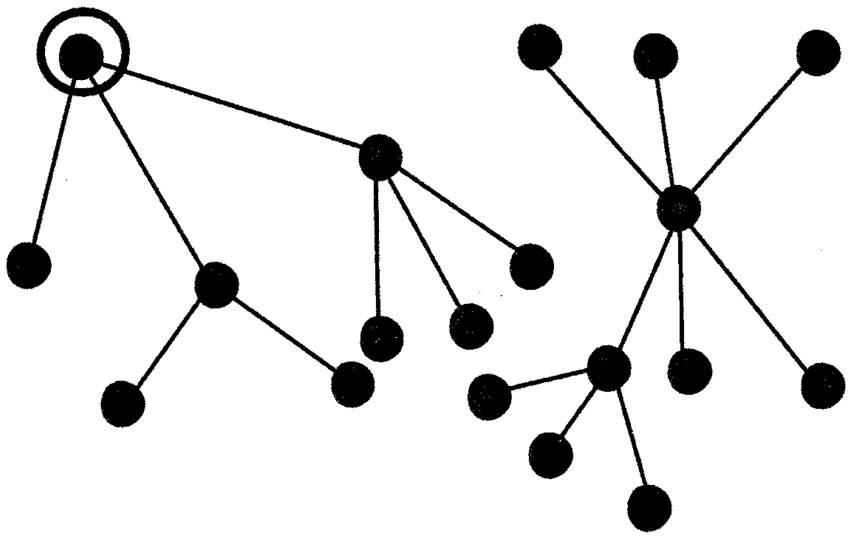
ここでの Network=計算機ネットワークとは，待ち行列ネットワークを指すものとする．待ち行列ネットワークの確率的解析に関しては，フィードバックがあるタイプも含め様々なネットワーク構造の解析がすでに行われている．特に定常状態の解が積形式解の形になる場合と条件について深く研究が進められている．

ここではネットワークの客の振り分け制御に確率過程ゲームを用いる．確率過程ゲームは，非協力ゲームにおける繰り返しゲームを精緻化したものである．それ故，プレイヤーの立場，戦略といった設定を明確にする必要がある．しかし，クローズしたネットワークの中のフロー（客の流れ）や，部分的にオープンであってもフィードバックのある場合はそのネットワークの内部のコントロールする主体の設定が困難であるためここでは，木構造に限った分析を行う．

2．対象とする待ち行列ネットワークに関する考察

2．1　木構造ネットワーク

本稿で対象とするのは木構造ネットワークである．ただし，下記の右図のような場合を考慮しない．頂点（根）から要求(客)が出され，それが階層的な構造を経て流れていくような場合を対象とする．つまり，下記の左図のような根つき木（二重丸になっているノードを根とする）の階層構造になるものを対象とする．(勿論，位相幾何学的には右図もあるノードを根として，根つき木として書き換えることはできるが，すべてのノードが同等なものは取り扱わない．)



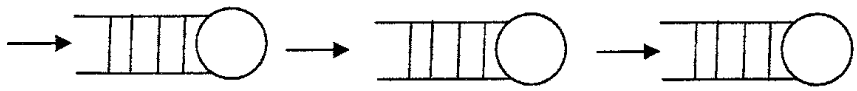
2．2　木構造の待ち行列ネットワーク

ここでは，上記で考えるノードは待ち行列ノードであり．階層を経るにつれ下位層に複数のノードがあれば，客の振り分け制御を確率過程ゲームで行う．ここでは，振り分け制御の結果ネットワーク全体が定常状態となる場合のみの考察を行う．待ち行列の研究の立場からは，フィードバックがある部分的，もしくは全体的にクローズされたネットワークが興味深い．しかし，制御問題になると困難を伴うためあえてこのような単純なネットワークを対象とする．

2．3　木構造待ち行列から直列待ち行列への分解

木構造待ち行列ネットワークで根と根以外のあるノードを特定する．すると，根からそのノードまではグラフが木であるためパスが唯一に決定される．したがってその間は直列待ち行列 (Tandem Queues) となる．また，その特別なノードを含みそれより下の階層の部分グラフにおいてまた別のノードを指定すれば同様に直列待ち行列である．

ここでは，直列待ち行列の重要な性質について，レビューをする．



このような直列待ち行列では，最初の待ち行列の出力過程が，第二の待ち行列の入力過程になる．解析の為，以下のような仮定をおく。すべて最初の待ち行列に関するものである．

- 入力過程はパラメータ λ のポアソン過程
- サービスの確率過程はパラメータ μ の指数分布にしたがう
- $\rho := \lambda / \mu$
- X_a ：到着間隔を示す確率変数
- X_S ：サービス間隔を示す確率変数
- X_d ：出力間隔を示す確率変数
- Y_q ：待ち行列にいる顧客数を示す確率変数
- 上記の確率変数に対しては，すべて確率密度関数が存在しそれは $f_X(\cdot)$ のように確率変数を添え字とする記号で表すものとする．

これはいわゆる M/M/1 でもっとも単純な待ち行列である．さて，2 番目の待ち行列の入力過程は最初の待ち行列の出力過程であることは明らかである。以下その為の解析を行う．

最初の待ち行列から客が出てきたとき，「次の客がすでに待ち行列に待機しておりすぐにサービスされる場合」(Case 1) と，「次の客が待ち行列にまだいない場合」(Case 2) があり，次の２つの式が得られる．

(Case 1)

$$P(X_d \le t \mid Y_q \ne 0) = P(X_s \le t)$$

(Case 2)

$$P(X_d \le t \mid Y_q = 0) = P(X_a + X_s \le t)$$

ここで，関数 f のラプラス変換が存在するとき，

$L(f)(s)$ ：f のラプラス変換,((s)は省略することもある)

と書くことにする．ラプラス変換の線形性を利用し，(Case1) の待ち行列に客がいる確率は ρ で，すると (Case2) の場合は $(1-\rho)$ の確率で起きることは明らかなので，以下の式が成り立つ．

$$L(fx_d)=(1-\rho)L(fx_a)L(fx_s)+\rho L(fx_s) \text{-----} (*)$$

指数分布のラプラス変換より

$L(fx_a)=\dfrac{\lambda}{\lambda+s}$　$L(fx_s)=\dfrac{\mu}{\mu+s}$　また定義より $\rho=\dfrac{\lambda}{\mu}$

これらを (*) に代入して計算する．$(*)=\dfrac{\lambda}{\lambda+s}$ が，単純計算によって得られる．この逆ラプラス変換は指数分布であるので，最初の待ち行列の出力過程はポアソン過程である．2 番目の待ち行列の入力過程もパラメータが同じポアソン過程となる。直列待ち行列ネットワークにおける最初の入力過程がポアソン過程であれば，指数サービスの待ち行列が連なっている場合，どの待ち行列においても入力仮定は同じパラメータのポアソン過程となることがわかる．

2．3　木構造待ち行列再考

木構造待ち行列を前節の考察をもとに再考することにする．本稿では客の振り分けについて確率過程ゲームによる制御を行うがネットワークの全体としては，定常状態におかれていることを仮定している．

この仮定によって，ある待ち行列に対して一階層下の待ち行列が k 個あるときに，制御の結果 1 番目から k 番目に振り分けられる確率分布 $(p_1,\ldots,p_i,\ldots,p_m)$ が存在する．すると，i 番目の待ち行列に対する入力は，元のポアソン過程の指数分布に p_i を掛けたものに他ならない．

3. 確率過程ゲーム（Stochastic Game）についての予備知識

確率過程ゲーム(Stochatic Game)は，ゲーム理論の標準的テキストに解説が記述されていることはほとんど無いため，ここで必要事項をMertens(2002)に従って記す．

3．1 確率過程ゲームの為の定義

- 状態空間；可測空間 $(\Omega, \mathcal{A})$
- プレイヤー； I （Iは可算有限集合とする）
- プレイヤー $i \in I$ のアクション空間；可測空間 $(S^i, \mathcal{S}^i)$
- 全アクション集合；

$$S = \prod_{i \in I} S^i$$

- 推移確率；

$$P : \Omega \times S \to \Omega$$

$$P(A|\omega, s) \text{ where } A \in \mathcal{A}, \omega \in \Omega, s \in S$$

- 利得関数ベクトル；$g(\omega, s)$
  これは下記で述べる各プレイヤーの利得関数 $g^i(\omega, s)$ $i \in I$ から構成されるベクトルである．
- ゲームの初期状態；ω_1
- ゲームの時間n（ここで離散時間で考える）の状態； ω_n
- プレイヤーはアクションベクトル； s_n（時間 n の各プレイヤーのアクションなのでベクトルであることに注意）を同時に選択する．そのとき各プレイヤーi は，$g^i(\omega_n, s_n)$を利得として受け取り，遷移確率に従って移る．
- この章では，上付き添え字がプレイヤーになっており，下付き添え字が時間に対応している．後に g_nが現れるがこれは時間nでの利得関数ベクトルの意味である．

3．2 割引率と確率過程ゲームの種類

確率過程ゲームは繰り返しゲームの延長線上にあるものと考えられるから，割引率をどのように考慮するかによって次のようなバリエーションがある．

- 割引ゲーム；Γ_λ
  - 各ステージの利得ベクトルに対して一期ごとに割引率λを考慮

- nステージゲーム；Γ_n
  - 時間nまでの時間平均利得ベクトルを考慮

- 割引なし（定常もしく平衡状態）ゲーム；Γ_∞
  - 上記の∞ステージの平均利得ベクトル $g_\infty = \lim_{n\to\infty} g_n$ を考慮

3．3 Historyと戦略

ここで使うHistoryは確率過程論におけるHistoryと厳密には異なる。確率過程論では状態の変化にのみ注目するがここではアクションという概念を導入されていることに注意すべきである．また，さらに注意を要するのは「戦略」である。確率過程ゲームは，Markov 決定過程論と深くかかわっている。「アクション」と「戦略」は非常に混同しやすいものである．

- プレイヤーiのHistoryとは
  - 状態とアクションの列$(\omega_1^i, s_1^i, \ldots, \omega_{n-1}^i, s_{n-1}^i, \omega_n^i)$

- プレイヤーiの戦略とは
  - $(\omega_1^i, s_1^i, \ldots, \omega_{n-1}^i, s_{n-1}^i, \omega_n^i) \to S^i$ の推移確率
  - ＜注意＞ゲーム理論の「混合戦略」になっている！

- Markov 戦略
  - Historyによらずにω_nにのみdependする

Markov 性は確率過程論では独立性を弱めた重要な概念である．次にどのような状態移行するか考えるときにすべてのHistoryを意思決定者が考慮するとはアプリケーションを考える上からも考えにくい．次の状態に移行するには現在の状態のみ依存するという考え方は理想的な単純化の一つである。多重Markov性を考える応用もありうるがそれは時間と状態をとりなおしてMarkovと考えられることも多い．

3．4 2人ゼロ和ゲームでの例

以下のような２人ゼロ和ゲームを考える

- ２つの状態Ω=｛1，2｝
- 利得行列はω=1 の時は左，ω=2 の時右であるとする（利得行列の値はプレイヤー１を示すものとする.）

$$
\begin{bmatrix} 10 & -1 \\ -1 & 10 \end{bmatrix} \quad \begin{bmatrix} 0 & 6 \\ 6 & 0 \end{bmatrix}
$$

- 遷移確率：ω=1 の時，（ここでは各行、各列はアクションを示し、行列要素の左が状態１へ、右が状態２への遷移確率を示している）

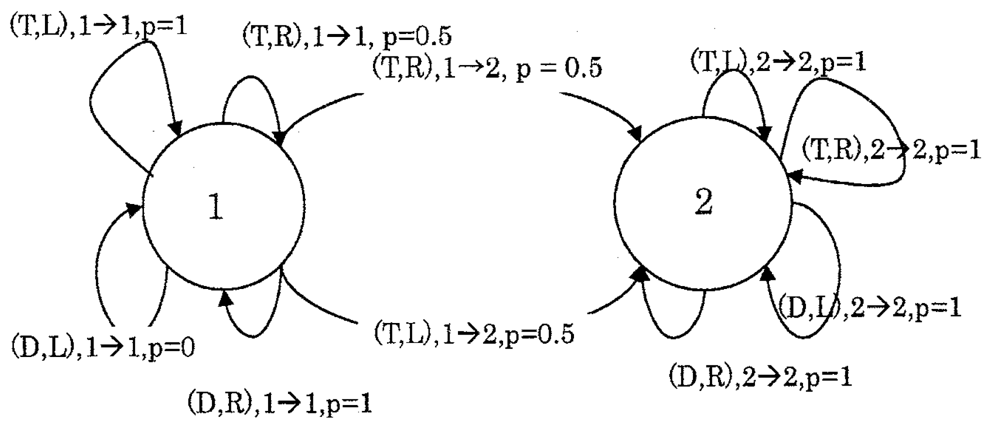
$$
\begin{pmatrix} (1,0) & (0.5,0.5) \\ (0.5,0.5) & (1,0) \end{pmatrix}
$$

- 遷移確率：ω=2 の時

$$
\begin{pmatrix} (0,1) & (0,1) \\ (0,1) & (0,1) \end{pmatrix}
$$

状態遷移図を描くと以下の通り，プレイヤー１の２つのアクションをT,Dとし，プレイヤー２を同様にL,Rとした．状態→状態で遷移を，pは遷移確率を示す．

状態遷移図



3．5　例に対する割引ゲームの解法

- v ： ω=1 の割引ゲーム値
- u ： ω=2 の割引ゲーム値
- 以下の２つの再帰式を計算（ただし，val は行列の min-max 値，割引率λ=0.5 とする）

$$v = val\begin{pmatrix} 0,5*10+0.5(1*v+0*u) & 0.5*(-1)+(0.5*v+0.5*u) \\ 0.5*(-1)+(0.5*v+0.5*u) & 0,5*10+0.5(1*v+0*u) \end{pmatrix}$$

$$u = val\begin{pmatrix} 0,5*0+0.5*u & 0.5*6+0.5*u \\ 0,5*3+0.5*u & 0.5*0+0.5*u \end{pmatrix}$$

$$u = 0.5*u+0.5val\begin{pmatrix} 0 & 6 \\ 6 & 0 \end{pmatrix} = 0.5u+1$$

これより u=2 同様にして v=4 がわかる．

最適戦略は以下のようになることがわかる．
ω=1 の時：
　　プレイヤー１ (0.5, 0.5)，プレイヤー２ (0.5, 0.5)
ω=2 の時：
　　プレイヤー１ (1/3, 2/3)，列プレイヤー２ (2/3, 1/3)

4．木構造ネットワークへの確率過程ゲームの適用

4．1　モデル

ある階層の待ち行列が，一階層下のｋ個の並列待ち行列に対して客を割り当てる状況を考える．Mohri(2005)と議論とほぼ同様の議論ができる．

- 非協力ゲームとしてのプレイヤーは，ルータとｋ個待ち行列の集合体のサーバの２プレイヤーとする．（待ち行列の集合体をまとめてプレイヤーと考えていることに注意する）
- 到着に関する考察は２章の結果を利用．
- 各待ち行列 のバッファは無限大であると仮定する．
- 各待ち行列の処理方式はFCFSであるとする．
- 下層の待ち行列のサービス率は各待ち行列 i に対して指数サービスを行なうものとし，そのパラメータは $a_1(i) \in [\underline{\mu}_i, \overline{\mu}_i]$ とする（下付き添え字の１はプレイヤー１を示している），これがプレイヤー１のアクションである．
- プレイヤー２がどの待ち行列に客を割り当てるかというアクションを a_2 で表現して，そのとる値は割り当てられる待ち行列の番号であるとする．

プレイヤー１と２の間には非協力ゲーム状況があり，各待ち行列のサービス率がルータには厳密には分からないとする．ここでは以下の情報が分かっていると仮定する．

– 各時点の待ち行列 の長さ
– ルータがどういった振り分け履歴があるかは分かっている

この問題をマルコフ決定過程として考えたい時に

- 状態：各待ち行列の長さ、直近のイベントの同定

が必要になる．ここで注意したいのは，ルータが「客が到着時に意思決定する」のに対して，サーバは「いつでも意思決定する」ことである．この状況を単純化するため，ルータに関して「すべてのイベントで意思決定」可能と考えるとモデルを単純化できる．つまり，

ルータが時間ｔでアクションｉを取る

↕　読替え

時間ｔ以降直近のイベントが到着なら客を待ち行列 ｉ に振り分ける

・状態：各待ち行列の長さ（だけでよい）

ということになる．この状況で具体的に定式化を試みる．

$A_i(s)$: 状態sでqueue iへの客が到着した直後の状態

$D_i(s)$: 状態sでqueue iへの客が出発した直後の状態

微小時間間隔：Δ　,$\pi\{p\}$: pが真なら１、偽なら０を返す

状態の変化（各$queue$の長さの変化）$s=(s_1,s_2,\ldots,s_k)\to s'$

遷移率（確率ではないことに注意）は以下のようになる

$$q_\Delta(s'\mid s,a_1,a_2)=\begin{cases}\lambda' p_i\Delta+o(\Delta) & (a_2=i,s'=A_i(s),i=1,2,\ldots,k)\\ a_1(i)\Delta+o(\Delta) & (s'=D_i(s),s_i\neq 0,i=1,2,\ldots,k)\\ p_i-(\lambda' p_i+\sum_{i=1}^{k}a_1(i)\pi\{s_i>0\})+o(\Delta) & (s'=s)\\ o(\Delta) & (\text{otherwise})\end{cases}$$

このように議論を Mohri(2005)と同様に展開することができる．

４．２　最適戦略について

解析を進めると推移率、利得関数とも

– プレイヤー１部分+ プレイヤー２部分

　となっている

このような両方の性質を満たすことを <u>AR-AT 性</u>(Additive Reward-Additive Transition) と呼び Raghavan,Tijs and Vrieze(1985)では，AR-AT 性を満たすなら以下の存在定理を証明し，それをもとめるアルゴリズムを示している．

- 有限計画期間の場合、Markov 最適戦略が存在
- 無限計画期間の場合、定常戦略の存在

本稿での前提はて定常性の仮定をおいているために有限期間に対しては,

この定理が無限計画期間利得（コスト）と最適政策（戦略）を求めることはできない．これにはある意味での過渡解析が必要になる。無限期間にのみこの定理が有効である．

5． まとめと今後の課題

本論では，木構造の待ち行列ネットワークに対する確率過程ゲームを用いた振り分けについて考察を行った．定常性を最初から仮定している点は改善されるべき一番の点である．これには，過渡解析が必要である．また，フィードバックを部分的に含んだ構造，特に直列待ち行列にフィードバックの部分があるような場合について解析を行うことが今後の課題である．

[参考文献]

[1]E. Altman. "A Markov game approach for optimal routing into a queuing network", 1994, Technical Report 2178, INRIA Sophia Antipolis.

[2]J. Filar and K. Vrieze, "Competitive Markov Decision Processes",1996, Springer.

[3]R. Hassin and M. Haviv, "To Queue or Not To Queue", 2004, Kluwer Academic Press.

[4]J.F. Mertens, "Stochastic Games", Handbook of Game Theory with Economic Application Vol.3, 2002, Chapter 47, pp.1809-1832.

[5]T.E.S. Raghavan and J.A. Filar, "Algorithms for Stochastic Games: -A Survey", ZOR(Mathematical Methods of OR) , 1991, pp.437-472

[6]L.S. Shapley, "Stochastic Games", Proceedings Nat. Acad. Of Science USA, 1953, 39, pp.1095-1100.

[7]S.H. Tijs, T.E.S. Raghavan and O.J. Vrieze, "On Stochastic Games with Additive Reward and Transition Structure", Journal of Optimization Theory and Applications, 1985, 47, pp.451-464.

[8]N. Vieille, "Stochastic Games: Recent Results", Handbook of Game Theory with Economic Application Vol.3, 2002, Chapter 48, pp.1833-1850.

[9]H. Mohri, "確率過程ゲームの計算機ネットワーク制御への応用"，京都大学数理解析研究所 講究録 (to appear)